

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

IN RE APPLICATION OF: Koichi MASUKURA, et al.

GAU:

SERIAL NO: New Application

EXAMINER:

FILED: Herewith

FOR: MOVING-PICTURE PROCESSING METHOD AND MOVING-PICTURE PROCESSING APPARATUS

REQUEST FOR PRIORITY

COMMISSIONER FOR PATENTS  
ALEXANDRIA, VIRGINIA 22313

SIR:

- ☐ Full benefit of the filing date of U.S. Application Serial Number , filed , is claimed pursuant to the provisions of 35 U.S.C. §120.
- ☐ Full benefit of the filing date(s) of U.S. Provisional Application(s) is claimed pursuant to the provisions of 35 U.S.C. §119(e): Application No. Date Filed

- ☒ Applicants claim any right to priority from any earlier filed applications to which they may be entitled pursuant to the provisions of 35 U.S.C. §119, as noted below.

In the matter of the above-identified application for patent, notice is hereby given that the applicants claim as priority:

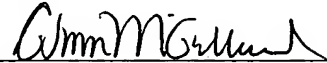
<u>COUNTRY</u>	<u>APPLICATION NUMBER</u>	<u>MONTH/DAY/YEAR</u>
Japan	2002-332756	November 15, 2002

Certified copies of the corresponding Convention Application(s)

- ☒ are submitted herewith
- ☐ will be submitted prior to payment of the Final Fee
- ☐ were filed in prior application Serial No. filed
- ☐ were submitted to the International Bureau in PCT Application Number  
Receipt of the certified copies by the International Bureau in a timely manner under PCT Rule 17.1(a) has been acknowledged as evidenced by the attached PCT/IB/304.
- ☐ (A) Application Serial No.(s) were filed in prior application Serial No. filed ; and
- ☐ (B) Application Serial No.(s)  
☐ are submitted herewith  
☐ will be submitted prior to payment of the Final Fee

Respectfully Submitted,

OBLON, SPIVAK, McCLELLAND,  
MAIER & NEUSTADT, P.C.



Marvin J. Spivak

Registration No. 24,913

C. Irvin McClelland  
Registration Number 21,124

Customer Number

22850

Tel. (703) 413-3000  
Fax. (703) 413-2220  
(OSMMN 05/03)

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日            2 0 0 2 年 1 1 月 1 5 日  
Date of Application:

出 願 番 号            特 願 2 0 0 2 - 3 3 2 7 5 6  
Application Number:

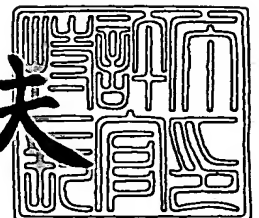
[ST. 10/C]:            [ J P 2 0 0 2 - 3 3 2 7 5 6 ]

出      願      人            株 式 会 社 東 芝  
Applicant(s):

2 0 0 3 年    7 月 1 8 日

特許庁長官  
Commissioner,  
Japan Patent Office

今 井 康 夫



【書類名】 特許願

【整理番号】 A000205327

【提出日】 平成14年11月15日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 15/00

【発明の名称】 動画像加工方法及び装置

【請求項の数】 12

【発明者】

【住所又は居所】 神奈川県川崎市幸区小向東芝町 1 番地 株式会社東芝研究開発センター内

【氏名】 増倉 孝一

【発明者】

【住所又は居所】 神奈川県川崎市幸区小向東芝町 1 番地 株式会社東芝研究開発センター内

【氏名】 堀 修

【発明者】

【住所又は居所】 神奈川県川崎市幸区小向東芝町 1 番地 株式会社東芝研究開発センター内

【氏名】 金子 敏充

【発明者】

【住所又は居所】 神奈川県川崎市幸区小向東芝町 1 番地 株式会社東芝研究開発センター内

【氏名】 三田 雄志

【発明者】

【住所又は居所】 神奈川県川崎市幸区小向東芝町 1 番地 株式会社東芝研究開発センター内

【氏名】 山本 晃司

**【発明者】**

**【住所又は居所】** 神奈川県川崎市幸区小向東芝町 1 番地 株式会社東芝研究開発センター内

**【氏名】** 大盛 善啓

**【特許出願人】**

**【識別番号】** 000003078

**【氏名又は名称】** 株式会社 東芝

**【代理人】**

**【識別番号】** 100058479

**【弁理士】**

**【氏名又は名称】** 鈴江 武彦

**【電話番号】** 03-3502-3181

**【選任した代理人】**

**【識別番号】** 100084618

**【弁理士】**

**【氏名又は名称】** 村松 貞男

**【選任した代理人】**

**【識別番号】** 100068814

**【弁理士】**

**【氏名又は名称】** 坪井 淳

**【選任した代理人】**

**【識別番号】** 100092196

**【弁理士】**

**【氏名又は名称】** 橋本 良郎

**【選任した代理人】**

**【識別番号】** 100091351

**【弁理士】**

**【氏名又は名称】** 河野 哲

## 【選任した代理人】

【識別番号】 100088683

【弁理士】

【氏名又は名称】 中村 誠

## 【選任した代理人】

【識別番号】 100070437

【弁理士】

【氏名又は名称】 河井 将次

## 【手数料の表示】

【予納台帳番号】 011567

【納付金額】 21,000円

## 【その他】

国等の委託研究の成果に係る特許出願（平成14年度通信・放送機構「ユーザーオリエンテッドマルチメディア技術の研究開発」委託研究、産業活力再生特別措置法第30条の適用を受けるもの）

## 【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書  
【発明の名称】 動画像加工方法及び装置  
【特許請求の範囲】

【請求項 1】 複数の時空間領域を有する入力動画像について、各々の時空間領域に関する情報を含むメタデータを取得する取得ステップと、

前記複数の時空間領域の少なくともいずれかに対応する切り出し領域を前記メタデータに基づいて決定するステップと、

前記入力動画像の各フレームの画像から前記切り出し領域を切り出して加工するステップと

を具備することを特徴とする動画像加工方法。

【請求項 2】 前記切り出し領域を、前記メタデータに含まれる画像特徴量、音声特徴量、意味特徴量の少なくともいずれかに関する情報に基づいて決定するステップを具備することを特徴とする請求項 1 に記載の動画像加工方法。

【請求項 3】 前記切り出し領域を前記メタデータに含まれる出力動画像の使用者、使用機器、使用回線、使用目的、課金情報の少なくともいずれかを表す情報に基づいて決定するステップを具備することを特徴とする請求項 1 又は 2 に記載の動画像加工方法。

【請求項 4】 前記切り出し領域を、前記メタデータに含まれる所定の切り出し領域位置制限又はカメラワークのパラメータ列の少なくともいずれかを表す情報に基づいて制限するステップを具備することを特徴とする請求項 1 乃至 3 のいずれかに記載の動画像加工方法。

【請求項 5】 少なくとも一つのフレームについて予め決定された切り出し領域に基づいて、他のフレームの切り出し領域を決定するステップを具備することを特徴とする請求項 1 乃至 4 のいずれかに記載の動画像加工方法。

【請求項 6】 前記メタデータに基づいて、前記切り出し領域の画像に表示画像を追加し、又はフィルタ処理するステップをさらに具備することを特徴とする請求項 1 乃至 5 のいずれかに記載の動画像加工方法。

【請求項 7】 複数の時空間領域を有する入力動画像について、各々の時空間領域に関する情報を含むメタデータを取得する手段と、

前記複数の時空間領域の少なくともいずれかに対応する切り出し領域を前記メタデータに基づいて決定する手段と、

前記入力動画像の各フレームの画像から前記切り出し領域を切り出して加工する手段と

を具備することを特徴とする動画像加工装置。

【請求項 8】 前記切り出し領域を、前記メタデータに含まれる画像特徴量、音声特徴量、意味特徴量の少なくともいずれかに関する情報に基づいて決定する手段を具備することを特徴とする請求項 7 に記載の動画像加工装置。

【請求項 9】 前記切り出し領域を前記メタデータに含まれる出力動画像の使用者、使用機器、使用回線、使用目的、課金情報の少なくともいずれかを表す情報に基づいて決定する手段を具備することを特徴とする請求項 7 又は 8 に記載の動画像加工装置。

【請求項 10】 前記切り出し領域を、前記メタデータに含まれる所定の切り出し領域位置制限又はカメラワークのパラメータ列の少なくともいずれかを表す情報に基づいて制限する手段を具備することを特徴とする請求項 7 乃至 9 のいずれかに記載の動画像加工装置。

【請求項 11】 少なくとも一つのフレームについて予め決定された切り出し領域に基づいて、他のフレームの切り出し領域を決定する手段を具備することを特徴とする請求項 7 乃至 10 のいずれかに記載の動画像加工装置。

【請求項 12】 前記メタデータに基づいて、前記切り出し領域の画像に表示画像を追加し、又はフィルタ処理する手段をさらに具備することを特徴とする請求項 7 乃至 11 のいずれかに記載の動画像加工装置。

#### 【発明の詳細な説明】

#### 【0001】

#### 【発明の属する技術分野】

本発明は、入力動画像に付帯する特徴量、出力動画像使用方法、切り出し位置制御情報等の各種付帯情報（メタデータ）を利用して、入力動画像の各フレームの画像から任意の矩形領域を切り出して加工することにより出力動画像を作成する動画像加工方法及び装置に関する。

## 【0002】

## 【従来の技術】

近年、画像処理技術の急速な発展により、動画像や静止画像をデジタルデータの形態で扱うことは一般化している。画像のデジタル化は、データ量が大きい動画像等の画像データを効率的に圧縮する技術を確立させた。また、このような技術の向上に伴う携帯電話機や携帯情報端末などの携帯型各種電子機器（「携帯機器」という）の急速な普及は、携帯機器上でも動画像を閲覧したいという一般的なユーザからの要望を生み出している。

## 【0003】

携帯機器は接続回線容量が小さくディスプレイの解像度や記憶容量も小さいことから、快適な閲覧のためには携帯端末用の動画像を別途作成する必要がある。このような携帯端末用の動画像を得るにあたっては、既存の動画像を国際標準であるMPEG-4などの携帯機器向けの動画像符号化フォーマットに効率的に変換する手法が既に提案されている（例えば下記非特許文献1参照。）。

## 【0004】

また、動画像利用の利便性向上ならびに簡易化を目的として、動画像の物理的特徴や意味的情報などの付帯情報（メタデータ）に応じた動画像の検索、編集、配信、閲覧等を実現する際の統一的な枠組みが求められており、メタデータに関する国際標準の一つにMPEG-7が提案されている。MPEG-7では、動画像や音声の物理的特徴量、内容などの意味的特徴量、著作権情報などを組み合わせて記述することができる。今後は、このようなMPEG-7の規格に則り、メタデータと共に動画像を扱う局面が急増するものと思われる。

## 【0005】

## 【非特許文献1】

山口 昇, 児玉 知也, 増倉孝一, MPEGトランスコード技術, 東芝レビュー, 57, 6, 2002, p18-21

## 【0006】

## 【発明が解決しようとする課題】

配信等に供するために携帯機器用の動画像を準備する際には、例えばその機種



ごとで異なる画面解像度や記憶容量等に応じ、それぞれ異なる動画像を作成する必要があることから、作業に極めて手間がかかる。

#### 【0007】

また、一般に携帯機器による映像再生には解像度が低い、画面が小さい、あるいは画面のアスペクト比が縦長であるといった特徴がある。したがって、携帯機器用の動画像とするために元々はテレビジョンやパーソナルコンピュータでの再生が意図された既存の映像をそのまま携帯機器の解像度に拡大あるいは縮小変換すると、アスペクト比の不具合や、小さい物体や小さい文字などが判別不能になるといった欠点が生じるかもしれない問題がある。

#### 【0008】

本発明はかかる事情を考慮してなされたものであり、動画像に付帯するメタデータを利用することにより入力動画像を適切に加工して出力動画像を得ることができる動画像加工方法及び装置を提供することを目的とする。より具体的には、入力動画像を構成するフレームの画像からの領域の切り出しをメタデータに基づいて適切に行うことである。

#### 【0009】

##### 【課題を解決するための手段】

本発明は入力動画像を構成する各フレームの画像から任意の領域を切り出して加工するものであり、その結果として得られた画像から構成される動画像を出力動画像とする。より具体的には、入力動画像に付帯するメタデータ中に含まれる任意の時空間領域に関する情報を利用し、少なくとも1つの時空間領域について、その少なくとも一部が出力動画像に含まれるように切り出し領域を決定するものである。ここでいう時空間領域とは、入力動画像の少なくとも一部の画像特徴に基づいて抽出される領域のことであり、時間的および空間的な広がりを持つ一塊の領域に相当する。入力動画像としては、原画像データそのもの、あるいは予め符号化されているデータを含む。また、入力動画像を構成する各フレームの画像からの切り出し領域は、矩形領域を含む。

#### 【0010】

時空間領域に関する複数の情報に基づいて、ある時空間領域については出力動

画像に含め、他のある時空間領域については出力動画像に含めないように切り出し領域を決定してもよい。

#### 【0011】

また、付帯するメタデータにおいて示される入力動画像の色、動き、テクスチャ、カット、特殊効果、物体の位置、文字情報などの画像特徴量、音の大きさ、周波数スペクトラム、波形、発話内容、音色などの音声特徴量、場所、時間、人物、感情、イベント、重要度、リンク情報などの意味特徴量、出力動画像の使用者、使用機器、使用回線、使用目的、課金情報などの使用情報のいずれかを利用して切り出し領域を決定してもよい。

#### 【0012】

また、予め計算された複数フレームにわたる切り出し領域の位置、予め作成された切り出し位置の制限情報およびカメラワークのパラメータ列などの切り出し位置制御情報のいずれかを利用して切り出し領域を決定してもよい。

#### 【0013】

##### 【発明の実施の形態】

図1は、本発明の一実施形態に係る画像処理装置の構成を示すブロック図である。図1に示されるように、本装置は、入力動画像記憶部101と、メタデータ記憶部102と、切り出し領域決定部103と、動画像切り出し部104と、出力動画像表示部105とから構成されている。本実施形態は、例えば汎用の計算機（コンピュータ）および同計算機上で動作するソフトウェアとを用いて実現することができ、図1に示す構成要素の一部は、オペレーティングシステムの下で動作するコンピュータプログラムのモジュールとして実現できる。

#### 【0014】

入力動画像記憶部101は、入力する動画像もしくは動画像符号化データを記憶するものであり、例えばハードディスクや光ディスク、半導体メモリ等により構成される。入力動画像記憶部101は、入力する動画像もしくは動画像符号化データを出力できるものであればどのようなものでもよく、たとえばビデオカメラや放送波チューナーなどでもよい。

#### 【0015】

メタデータ記憶部102は、入力動画像の特徴量や出力動画像の使用方法、切り出し位置制御情報等の各種付帯情報（メタデータ）を記憶するものであり、メタデータ記憶部101と同様、ハードディスクや光ディスク、半導体メモリ等により構成される。メタデータを入力動画像のデータに対してどのように関連付け、付帯させるかについては任意である。たとえば、メタデータが複数に分割されていてもよいし、複数の物理デバイス上に存在していてもよい。また、入力動画像記憶部101に記憶されている入力動画像と一体化していてもよい。また、メタデータは、入力動画像を解析して取得してもよいし、出力機器やその出力回線を解析して取得してもよい。あるいは、処理中にユーザーが直接、メタデータを入力してもよい。

#### 【0016】

入力動画像記憶部101およびメタデータ記憶部102は、同一の物理デバイス上に存在していても、異なる物理デバイス上に存在していてもよい。あるいは、ネットワークや放送波を経由して遠隔地に存在していてもよい。

#### 【0017】

切り出し領域決定部103は、メタデータ記憶部102に記憶されているメタデータを読み込み、該メタデータに基づいて入力動画像を構成する各フレームの画像における切り出し領域（例えば矩形の領域）を決定する。基本的に、切り出し領域は1フレームごとに決定することとするが、複数フレームの切り出し領域を一度に決定したり、一度決定した切り出し領域を他のフレームの切り出し領域やメタデータに応じて変更するよう構成することもできる。

#### 【0018】

動画像切り出し部104は、切り出し領域決定部103で決定された切り出し領域情報に従って、入力動画像記憶部101に記憶されている入力動画像の各フレームの画像から、切り出し領域決定部103により決定された切り出し領域に相当する画像領域を切り出して出力動画像を作成する。この動画像切り出し部104において、切り出し操作を施す前の各フレームの画像もしくは切り出し操作を施した後の各フレームの画像に対し、拡大縮小、回転、フィルタリングなどの各種画像処理加工を施してもよい。また、出力動画像を例えば動画像符号化の国

際標準であるMPEG-1, 2, 4等に基づいて符号化し、動画像符号化データを作成してもよい。

#### 【0019】

出力動画像表示部105は、動画像切り出し部104で作成された出力動画像を表示するものであり、CRTや液晶など、動画像を表示（映像再生）できる画面を備えた機器であればどのようなものでも良い。例えば、パーソナルコンピュータ、携帯電話機、携帯情報端末などがあげられる。動画像切り出し部104により動画像符号化データを作成する構成とした場合は、該動画像符号化データを出力動画像表示部105が動画像に逆符号化してから表示する。出力動画像表示部105はネットワークや放送波を経由し、遠隔地に存在していてもよい。

#### 【0020】

出力動画像保存部106は、動画像切り出し部104で作成された出力動画像を保存するためのものであり、例えばハードディスクや光ディスク、半導体メモリ等により構成される。この出力動画像保存部106は、出力動画像表示部105と同様にネットワークや放送波を経由し遠隔地に存在していてもよい。

#### 【0021】

出力動画像表示部105および出力動画像保存部106は用途に応じて少なくともいずれか一方が必要である。もちろん、両者を備える構成としてもよい。

#### 【0022】

図2はメタデータのデータ構造の一例を示す図である。本例のメタデータは、入力動画像関連情報201、出力動画像使用情報202、切り出し位置制御情報203を有する。これら入力動画像関連情報201、出力動画像使用情報202、切り出し位置制御情報203のうち、少なくともいずれかの情報が必要である。また、それぞれの情報について、複数の情報項目が存在していてもよい。

#### 【0023】

また、入力動画像関連情報201、出力動画像使用情報202、切り出し位置制御情報203の配置構成（例えば順序が挙げられる）は任意である。例えば、出力動画像使用情報202内に切り出し位置制御情報203が含まれるような木構造としてもよい。つまり、メタデータのデータ構造は図2に示した以外のもの

も考えられるが、以下詳細に説明する所要の情報を格納し、必要に応じて読み出すことができるものであれば、どのような構造としてもよい。例えば、国際標準である M P E G - 7 を利用することもできる。

#### 【0024】

入力動画像関連情報 201 は、入力動画像もしくは入力動画像符号化データに関する時空間領域情報 211 および特徴量情報 212 を含む。時空間領域情報 211 および特徴量情報 212 のうち少なくとも一方が必要である。また、それぞれの情報について、複数の情報項目が存在していてもよい。

#### 【0025】

また、時空間領域情報 211 および特徴量情報 212 の配置構成（例えば順序）は任意である。例えば、時空間領域情報 211 内に特徴量情報 212 を記述するなど木構造としてもよい。この場合、時空間領域情報 211 に記載の時空間領域が特徴量情報 212 に記載の特徴量を持つこと、が木構造において表現される。

#### 【0026】

時空間領域情報 211 は、入力動画像の時間的および空間的な広がりを持つ一塊の領域を表すためのものであり、ヘッダ情報 221、開始終了時刻（データ）222、軌跡データ 223 を含む。開始終了時刻 222 および軌跡データ 223 のうち、少なくとも一方が必要である。また、それぞれのデータについて、複数のデータ項目が存在していてもよい。

#### 【0027】

ヘッダ情報 221 は、当該時空間領域情報の識別番号や名前を表し、および、開始終了時刻 222 および軌跡データ 223 のデータ形式を表す。

#### 【0028】

開始終了時刻 222 は、当該時空間領域の開始時刻および終了時刻を表す。開始終了時刻 222 は、時刻を一意に特定できるものであれば、どのような形式でもよい。例えば、入力動画像のタイムスタンプやフレーム番号、入力動画像撮影時の日時などが利用できる。

#### 【0029】

軌跡データ 223 は、時空間領域の形状を表現するためのパラメータである。先頭時刻から最終時刻までの時空間領域の形状を表現できるものであれば、軌跡データ 223 としてどのようなデータを用いてもよい。例えば、MPEG-7 の SpatioTemporalLocator 等により軌跡データ 223 を記述することができる。これは、各フレームの領域形状を矩形、楕円、多角形などにより表現するものであり、例えば、領域形状が矩形や多角形の場合は各頂点の軌跡を関数近似して得られるパラメータ、領域形状が楕円の場合は楕円の外接矩形頂点の軌跡を関数近似して得られるパラメータ等に相当する。

#### 【0030】

特徴量情報 212 は、ヘッダ情報 224、特徴量データ 225 から構成される。ヘッダ情報 224 は、当該特徴量がいかなる種類の特徴量であるかや、特徴量データ 225 がいかなるデータ形式で格納されているかを表す情報である。特徴量としては、ここでは画像特徴量、音声特徴量、あるいは意味特徴量を想定するが、入力動画像に関するものであれば、どのような特徴量を用いてもよい。

#### 【0031】

画像特徴量は、入力動画像の少なくとも 1 フレーム以上にわたる画像や画像列の一部もしくは全部についての色、動き、テクスチャ、カット、特殊効果、物体の位置、文字情報などの物理的特徴量や、既知の物理的特徴量から推定される特徴量の少なくともいずれかを含む。

#### 【0032】

音声特徴量は、入力動画像の少なくとも 1 つの音声チャンネルの一部もしくは全部について、音の大きさ、周波数スペクトラム、波形、発話内容、音色などの物理的特徴量や既知の物理的特徴量から推定される特徴量の少なくともいずれかを含む。

#### 【0033】

意味特徴量は、入力動画像の一部もしくは全部についての場所、時間、人物、感情、イベント、重要度、リンク情報などの動画像内容記述や、意味的な特徴量の少なくともいずれかを含む。

#### 【0034】

特徴量データ 225 は、当該特徴量情報に記された特徴量がいかなるものであるかを示す実際のデータであって、ヘッダ情報 224 中に指定された特徴量の種類に応じた所定のデータ形式に従って格納される。特徴量データ 225 は、例えば、色に関するものであればカラーヒストグラムで表現することができる。あるいは、場所に関するものであれば地名や緯度経度で表現することができる。このような特徴データ 225 の表現形式（データ形式）は、当該特徴量を特定できるものであれば、どのようなものでもよい。

#### 【0035】

出力動画像使用情報 202 は出力動画像の使用に関する情報を表すものであり、ヘッダ情報 231、使用情報データ 232 を含んでいる。なお、使用情報データ 232 内に出力動画像使用情報 202 が含まれるような木構造としてもよい。

#### 【0036】

ヘッダ情報 231 は、当該使用情報がいかなる種類の情報であるか、その具体的な使用情報データ 232 がいかなるデータ形式で格納されているかを表す。出力動画像使用情報 202 は、使用者や使用機器など、出力動画像の使用に関する情報であれば、どのようなものでもよい。

#### 【0037】

ここでいう使用者とは、出力動画像を使用する者のことであり、使用者を特定するための名前や ID、当該使用者がどのようなグループに含まれているかを表す情報、使用目的、課金情報などが出力動画像使用情報 202 に含まれる。

#### 【0038】

使用機器とは、出力動画像を閲覧する際に用いられる機器のことであり、機器名称、OS（オペレーティングシステム）、CPU 速度、画面解像度、サポートする動画像符号化形式、回線形式、回線速度などが出力動画像使用情報 202 に含まれる。

#### 【0039】

使用情報データ 232 は、当該使用情報の実際のデータであり、ヘッダ情報 231 に指定された使用情報の種類やデータ形式に従って格納される。データの格納方法は使用情報の種類に応じて異なり、例えば機器名称であれば文字列や ID

番号など当該使用情報を特定できるようなものであればどのようなものでもよい。

#### 【0040】

切り出し領域制御情報203は、切り出し領域の位置を制限するための情報や、カメラワークのパラメータ列などを規定するための情報である。カメラワークのパラメータ列としては、例えば、画像を拡大しすぎて画質が劣化することのないよう最大の可能拡大率を定めたり、切り出し領域が早く動きすぎることをないようにカメラワークに制約を課すためのパラメータなどに相当する。なお、出力動画画像がパンやズームなどのカメラワークを再現できるようにするためのカメラワークパラメータおよびその順序が記述されることもある。

#### 【0041】

図3は、本実施形態に係る画像処理装置が実行する処理手順の一例を示すフローチャートである。本処理手順は、メタデータ読み込みステップS31と、表示／非表示領域計算ステップS32と、切り出し領域計算ステップS33と、動画画像切り出しステップS34と、切り出し動画画像加工ステップS35と、動画画像出力ステップS36と、全フレーム終了判定ステップS37とにより構成されている。処理は、基本的には1フレームごとに行うが、全フレームを一度に行ってもよいし、数フレームごとなど複数フレームごとに行ってもよい。

#### 【0042】

先ずメタデータ読み込みステップS31において、メタデータ記憶部102からメタデータを読み込む。メタデータは、開始時に一度に全部読み込んでおいてもよいし、処理中に適宜読み込むようにしてもよい。次に表示／非表示領域計算ステップS32において、メタデータ内の時空間領域情報から当該フレームの表示領域と非表示領域を計算する。

#### 【0043】

ステップS32における表示／非表示領域計算の詳細を図4及び図5を参照して説明する。本処理によれば、例えば、視聴者が見たい重要な領域は不足なく含まれ、かつ不要な領域は含まれないように出力動画画像を作成することが可能になる。



## 【0044】

図4に示すように、入力動画像の当該フレームの画面401内に複数の時空間領域402、404が存在しており、時空間領域402は切り出す動画像に含める時空間領域とし、時空間領域404は切り出す動画像に含めない時空間領域とする場合を仮定する。

## 【0045】

ある時空間領域について、これが切り出す動画像に含める時空間領域であるか、切り出す動画像に含めない時空間領域であるかについては、メタデータの情報に基づいて区別することができる。その具体的な方法としては、時空間領域のヘッダ情報221に記述された識別番号や名前の条件で区別することができる。これには、例えば、識別番号や名前の先頭に区別する記号をあらかじめ付与するか、名前がある文字列に一致したら切り出す動画像に含めるようにすることなどが挙げられる。

## 【0046】

また、好ましくは、当該時空間領域の形状や軌跡データを利用して区別してもよい。例えば、形状が楕円のものは切り出す動画像に含めないとか、指定した点を通過する時空間領域は切り出す動画像に含めるなどの処理が行える。

## 【0047】

また、好ましくは、当該時空間領域情報内に記述されているメタデータや、メタデータ内に当該時空間領域情報が記述されている木構造の親ノードのメタデータや、当該時空間領域情報とメタデータにリンクが設定されているなど、当該時空間領域情報と関係する他のメタデータ内の情報を利用して区別してもよい。

## 【0048】

例えば、メタデータが当該時空間領域情報211内に記述された色や文字に関する画像特徴量が記述されているときは、赤い色をもつ時空間領域は切り出す動画像に含めるとか、テロップに相当する時空間領域であったなら切り出す動画像に含めないなどの処理が行える。また、出力動画像使用情報によって区別する方法を変化させると、ユーザーや使用機器に応じて時空間領域の処理を変えることができる。

## 【0049】

このように、切り出し動画像に含める時空間領域と含めない時空間領域が判別可能なときの1フレームの表示／非表示領域計算の一処理手順を図5に示す。この処理では、基本的に当該フレームに存在する時空間領域を一つずつ処理していくこととするが、複数の時空間領域を一度に処理してもよい。また、処理開始時には表示／非表示領域は存在しないものとするが、あらかじめ表示したくない領域が既知の場合などには、処理開始時に表示／非表示領域をあらかじめ設定しておいてもよい。

## 【0050】

時空間領域のフレーム形状取得ステップS61は、当該フレームにおける処理する時空間領域の形状を取得するためのステップである。時空間領域形状は通常、矩形や楕円や多角形で表されるため、それらの形状を表すパラメータを算出する。例えば、矩形や多角形であれば頂点座標列、楕円であれば外接長方形の頂点座標や長軸短軸の長さや回転角などとし、フレームの形状を一意に表せるものならばどのようなものでもよい。

## 【0051】

切り出し動画像に含める時空間領域であるかどうかをステップS62にて判別し、切り出し動画像に含める時空間領域であるならば、表示領域更新ステップS63にて表示領域を更新する。更新された表示領域は今までの表示領域と時空間領域のフレーム形状取得ステップS61で得られた領域のOR（論理和）領域のうち、画面401に含まれる部分となる。例えば、今までの表示領域が403であり、時空間領域のフレーム形状取得ステップS61によって得られた形状が402であるときは、領域403と402のOR領域のうち画面401に含まれる部分（図4において参照数字411が付与されたハッチング部分）となる。領域形状はパラメータによって表現されているため、表示領域は形状パラメータ列によって表すことができる。

## 【0052】

また、表示領域は、その形状や一度計算された表示／非表示領域に加工を施してもよい。例えば、時空間領域402の周辺に任意の余白406を追加してから

表示領域を計算したり、表示領域 411 を内包する最小の長方形（バウンディングボックス）412 を表示領域としてもよい。余白を追加するには、例えば時空間領域の重心座標を計算し、領域形状や外接矩形の各頂点と重心座標との距離が大きくなるように頂点座標を計算すればよい。

#### 【0053】

バウンディングボックスを表示領域とする場合は、従来の表示領域のバウンディングボックスが  $(X1, Y1) - (X2, Y2)$  で表され、ステップ S61 で得られる形状のバウンディングボックスが  $(x1, y1) - (x2, y2)$  で表され、画面 401 が  $(0, 0) - (W, H)$  で表されるとすると、更新された表示領域のバウンディングボックス 412 は  $(\max(0, \min(X1, x1)), \max(0, \min(Y1, y1))) - (\min(W, \max(X2, x2)), \min(H, \max(Y2, y2)))$  と表すことができ、このように簡易な計算で表示領域を算出できる。

#### 【0054】

切り出し動画像に含めない時空間領域であるかどうかをステップ S62 にて判別し、切り出し動画像に含めない時空間領域であるならば、非表示領域更新ステップ S65 にて表示領域を更新する。更新された非表示領域は今までの非表示領域と時空間領域のフレーム形状取得ステップ S61 で得られた領域の OR 領域のうち画面 401 に含まれる部分となる。例えば、今までの非表示領域が 405 であり、時空間領域のフレーム形状取得ステップ S61 によって得られた形状が 404 であるときは、領域 405 と 404 の OR 領域のうち画面 401 に含まれる部分（図 4 において参照数字 413 が付与されたハッチング部分）となる。領域形状はパラメータによって表現されているため、非表示領域は形状パラメータ列によって表すことができる。

#### 【0055】

また、ステップ S63 の処理と同様に、非表示領域は形状や一度計算された表示／非表示領域を加工して表示領域としてもよい。例えば、時空間領域周辺に余白を追加してから非表示領域を計算したり、非表示領域を内包する最小の長方形（バウンディングボックス）を非表示領域としてもよい。

#### 【0056】

全時空間領域完了判定ステップS66にて、当該フレームに存在するすべての全時空間領域について処理が終了したかどうかを判定し、全時空間領域について処理が終了するまで、ステップS61からステップS66までの処理手順を繰り返す。

#### 【0057】

図3に説明を戻す。切り出し領域計算ステップS33では、表示／非表示領域計算ステップS32で計算された表示／非表示領域とメタデータを利用して、当該フレームにおける入力動画像を切り出す領域を計算する。

#### 【0058】

ここで、図6及び図7を参照して、ステップS33における切り出し領域計算の詳細を説明する。

#### 【0059】

同図に示すように、入力動画像の当該フレームの画面501内に、表示領域502と非表示領域503とが存在する場合を仮定する。このとき、切り出し領域504は画面501内に内包され、表示領域502を内包し、非表示領域503と重なる部分がないような矩形領域であればどのような領域としてもよい。例えば、切り出し領域の重心を表示領域の重心と同じ位置にして、表示領域の全領域が切り出し領域にすべて含まれる最小の矩形領域となるように切り出し領域を決めるなどの方法がある。

#### 【0060】

図7は、切り出し領域計算の一処理手順を示すフローチャートである。

#### 【0061】

制限情報読み込みステップS71では、メタデータから当該フレームに関する制限情報を取得する。制限情報は出力動画像を使用する機器の画素数やアスペクト比、切り出し領域のX、Y方向それぞれの移動速度やその最大値、切り出し領域の最小の幅、高さ、面積、切り出し領域内の表示領域の位置関係など、切り出し領域の位置を制限するような情報である。制御情報はなくてもよいし複数存在していてもよい。

#### 【0062】

メタデータに直接、制限情報が記述されている以外にも、その他のメタデータやあらかじめ計算されているフレームの切り出し領域を利用して制限情報を生成してもよい。例えば、あらかじめ計算されているフレームの切り出し領域を利用して、計算するフレームの推定される切り出し領域位置を計算し、この推定された位置から一定距離以上切り出し領域が離れないように制限情報を生成すれば、切り出し領域がある方向に動いていたときに急に逆方向に動かないようにするなどの制御が可能である。また、切り出し領域の移動速度や加速度を一定以上にならないように制限情報を生成すれば、切り出し領域位置が振動しないようにすることなどが可能である。

#### 【0063】

メタデータの時空間領域情報を利用する場合は、例えば、入力画像領域を座標  $(0, 0) - (W, H)$  で囲まれる矩形、時空間領域中心を座標  $(x, y)$ 、切り出し領域を座標  $(X1, Y1) - (X2, Y2)$  で囲まれる矩形として、入力画像領域に対する時空間領域中心の相対位置と切り出し領域に対する時空間領域中心の相対位置が等しくなるようにしたり（すなわち、 $x/W = (x-X1)/(X2-X1)$ ,  $y/H = (y-Y1)/(Y2-Y1)$  になるように切り出し領域位置を制御する）、時空間領域の動きが速いところは切り出し領域を大きめに取るなどの制限情報を生成できる。

#### 【0064】

メタデータの色、動き、テクスチャ、カット、特殊効果、物体の位置、文字情報などの画像特徴量を利用する場合は、例えば、画面やそのなかに記述されている物体の動きベクトルやオプティカルフローを見て、動きが速い場面は切り出し領域を大きめに取ったり、物体の動きの進行方向に広く余白を取ることができる。また、カット情報が記述されているときは、カットとカットの間は切り出し領域があまりに急激に変化しないようにするなどの制限情報を生成できる。

#### 【0065】

メタデータの音の大きさ、周波数スペクトラム、波形、発話内容、音色などの音声特徴量を利用する場合には、例えば、発話内容が記述されているときは会話シーンにおいて話者を中心に切り出し領域を設定したり、音の大きさが記述されているときは音が小さいほど切り出し領域の時間的変化量を小さくして静かなシ

ーンとするなどの制限情報を生成できる。

#### 【0066】

メタデータの場所、時間、人物、感情、イベント、重要度、リンク情報などの意味特徴量を利用する場合は、例えば、野球のバッティングの際にはバッターを拡大するような切り出し領域にするなど、イベントごとに切り出し領域を調整したり、人物の感情の情報を見て落ち着いたシーンには切り出し領域の時間的変化量を小さくしたりするなどの制限情報を生成できる。

#### 【0067】

メタデータの使用者、使用機器、使用回線、使用目的、課金情報などの使用情報を利用する場合は、例えば、閲覧に使用する機器の画面の解像度情報から、切り出し後の1ピクセルの大きさが入力動画像の1ピクセルより小さくならないようにして画質の劣化を防いだり、使用者ごとに切り出し領域の中心とする物体を変化させたりするなどの制限情報を生成できる。

#### 【0068】

メタデータの切り出し位置の制限情報やカメラワークの順番などの切り出し領域制御情報を利用する場合は、例えば、切り出し領域制御情報に記載のカメラワークのパラメータ列と同様なカメラワークの出力動画像になるように切り出し領域を設定したり、切り出し領域の縦方向、横方向の時間的な動きの最大値が記述されているときは動きの最大値を超えないように切り出し領域を設定したり、できるだけ広い切り出し領域やできるだけ狭い切り出し領域になるように切り出し領域を設定したりするなどの制限情報を生成できる。

#### 【0069】

図7のフローにおいて、初期切り出し領域設定ステップS72では、切り出し領域の初期値を計算する。切り出し領域の初期値はどのように決めてもよいが、例えば一つ前のフレームで計算された切り出し領域を初期値としたり、表示領域のバウンディングボックスを初期値とするなどの方法がある。

#### 【0070】

次に、切り出し領域移動ステップS73では、表示／非表示領域や、ステップS71で読み込まれた制限情報に合致するように、切り出し領域位置を移動させ

る。表示／非表示領域や制限情報との合致度が高まる方向であれば、合致度の計算方法、移動方法、移動量は任意である。

#### 【0071】

例えば、切り出し領域が表示領域を内包していなかったら、切り出し領域外が表示領域の面積が減ると合致度が高まるように合致度を設定し、合致度が高まるように切り出し領域を拡大や移動させる。

#### 【0072】

切り出し領域505が $(X_{k1}, Y_{k1}) - (X_{k2}, Y_{k2})$ であって、表示領域502が $(X_{h1}, Y_{h1}) - (X_{h2}, Y_{h2})$ であるとする(ただし、 $X_{h1} < X_{k2} < X_{h2}$ ,  $Y_{k1} < Y_{h1}$ ,  $Y_{h2} < Y_{k2}$ )、切り出し領域外が表示領域の面積は $(X_{h2} - X_{k2}) * (Y_{k2} - Y_{k1})$ で計算されるので、 $X_{k2}$ を $X_{h2}$ 方向に動かすと切り出し領域外が表示領域の面積が減り、合致度を高めることができる。

#### 【0073】

制限情報と切り出し領域のアスペクト比が異なるときは、切り出し領域のアスペクト比と、制限情報のアスペクト比との割合が1になるほど合致度が高くなるような合致度を設定し、この合致度が高まるように幅や高さを拡大縮小する。すなわち、切り出し領域のアスペクト比を $\alpha_k (= (\text{幅}) / (\text{高さ}))$ 、制限情報のアスペクト比を $\alpha_s$ とすると、 $\alpha_k / \alpha_s$ が1に近づくほど合致度が高いため、 $\alpha_k / \alpha_s > 1$ のときは切り出し領域の幅を減らすか高さを増やす。逆に、 $\alpha_k / \alpha_s < 1$ のときは切り出し領域の高さを減らすか幅を増やす。

#### 【0074】

移動方法、移動量の決定方法としては、制限情報ごとにあらかじめ決めておいてもよいし、例えばニューラルネットワークなどの学習アルゴリズムを利用してもよい。

#### 【0075】

移動終了判定ステップS74では、切り出し領域が表示／非表示領域や制限情報に合致するかどうかを判別し、合致するまでステップS73を繰り返して切り出し領域を移動させる。表示／非表示領域や制限情報にすべて合致する切り出し領域が見つからないときは、適当な繰り返し回数で終了させてもよい。

**【0076】**

このように、メタデータの情報によって切り出し領域を調整することにより、入力動画像の内容や出力動画像の使用方法に適し、閲覧者が不自然に感じない出力動画像を作成することが可能となる。

**【0077】**

図3のフローにおいて、動画像切り出しステップS34では、以上説明した切り出し領域計算ステップS33で計算された切り出し領域を利用して、入力動画像のフレーム画像から、当該フレームの切り出し領域部分を切り出す。次に、切り出し動画像加工ステップS35では、動画像切り出しステップS34で作成された切り出し画像を加工等し、出力動画像を作成する。

**【0078】**

図8は、動画像切り出しの一処理手順を示すフローチャートである。図8に示されるように、本処理は画面拡大縮小回転ステップS81と画像加工処理ステップS82と動画像符号化ステップS83とから構成されている。画面拡大縮小回転ステップS81と画像加工処理ステップS82の処理順序は入れ替えてもよい。また、画面拡大縮小回転ステップS81、画像加工処理ステップS82、動画像符号化ステップS83は、それぞれの処理が必要ない際にはいずれかを省略してもよい。

**【0079】**

画面拡大縮小回転ステップS81では、動画像切り出しステップS34で作成された切り出し画像を拡大縮小したり回転させる。通常、切り出し画像の解像度はそれぞれ異なるが、動画像の解像度は一定でなければならないことが多いため、切り出し画像を動画像の解像度と等しくなるように拡大縮小を行う。また、閲覧に使用する機器によっては画像を90度回転させた映像のほうが閲覧しやすい場合がある。その場合には切り出し画像を90度回転させる。

**【0080】**

画像加工処理ステップS82では、メタデータの情報を利用して、切り出し画像のフィルタリングや、表示情報追加など各種の加工を行う。例えば、ある時空間領域の中や外にモザイクやぼかしなどのフィルタをかけたり、別の時空間領域



の画像を合成したり、文字情報や人物の名前などの情報をテロップとして画像中に表示させたりするなどの処理が可能である。もちろん、この場合にはメタデータを利用することができる。これら加工はいくつかを組み合わせてもよいし、その際の処理の順番も任意でよい。

#### 【0081】

動画像符号化ステップS83は、出力動画像を使用機器や使用回線に合わせて符号化データに圧縮するステップである。符号化フォーマットとしては国際標準であるMPEG-4などが通常使用されるが、用途に合わせてどのような符号化フォーマットでもよい。出力動画像を符号化する必要がないときは本ステップをスキップしてもよい。

#### 【0082】

そして、図3の動画像出力ステップS36では、切り出し動画像加工ステップS35で作成された出力動画像を用途に合わせて出力する。出力動画像を閲覧する際には使用機器側で再生表示が行われる。出力動画像を保存する際にはディスクやテープ等に保存される。出力動画像をネットワークや放送波で送信する際には、適する形式に変換し送信される。

#### 【0083】

次に、全フレーム終了判定ステップS37では、入力動画像の処理すべきフレームがすべて終了したかを判別する。全フレームが終了するまでステップS32からステップS37までの処理手順は繰り返される。

#### 【0084】

以上説明した本実施形態の画像処理装置によれば、入力動画像を構成するフレームの画像からの領域の切り出しをメタデータに基づいて行うことができ、入力動画像を適切に加工して出力動画像を得ることができる。これにより、例えば、配信等に応ずるために携帯機器用の動画像をその機種ごとで異なる画面解像度や記憶容量等に応じて準備するといった作業を容易に行うことができるようになる。また、解像度が低い、画面が小さい、あるいは画面のアスペクト比が縦長であるといった携帯機器の特徴については、メタデータを基に適切に画像加工を施すことにより、アスペクト比の不具合や、小さい物体や小さい文字などが判別不能

になるといった欠点が生じることもない。

#### 【0085】

なお、本発明は上述した実施形態に限定されず種々変形して実施可能である。

#### 【0086】

##### 【発明の効果】

以上説明したように、本発明によれば、メタデータの情報に応じて入力動画像から各フレームごとに適切な切り出し領域で切り出すことによって、内容や使用方法に応じた出力動画像を自動的に作成することが可能となり、閲覧する携帯端末にあわせた動画像を容易に作成できる。

##### 【図面の簡単な説明】

【図1】 本発明の一実施形態に係る画像処理装置の構成を示すブロック図

【図2】 メタデータのデータ構造の一例を示す図

【図3】 同実施形態に係る画像処理装置が実行する処理手順の一例を示すフローチャート

【図4】 表示／非表示領域計算を説明するための図

【図5】 表示／非表示領域計算の一処理手順を示すフローチャート

【図6】 切り出し領域の計算を説明するための図

【図7】 切り出し領域計算の一処理手順を示すフローチャート

【図8】 動画像切り出しの一処理手順を示すフローチャート

##### 【符号の説明】

101…入力動画像記憶部

102…メタデータ記憶部

103…切り出し領域決定部

104…動画像切り出し部

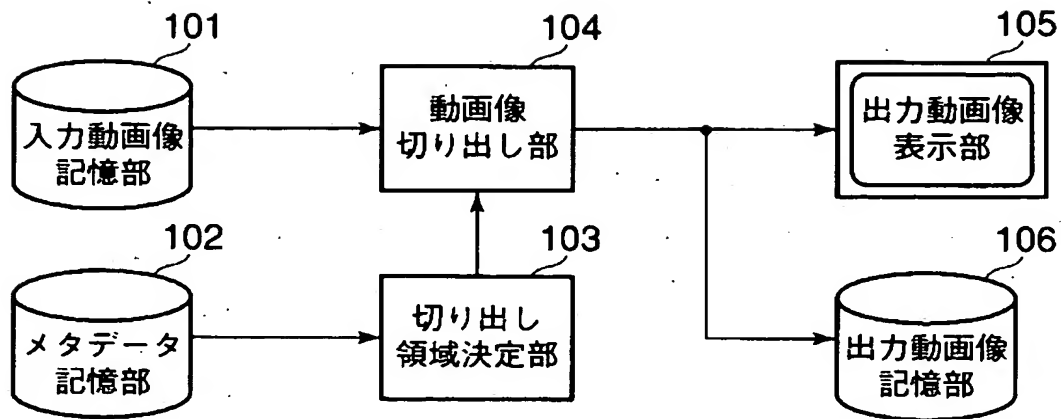
105…出力動画像表示部

106…出力動画像記憶部

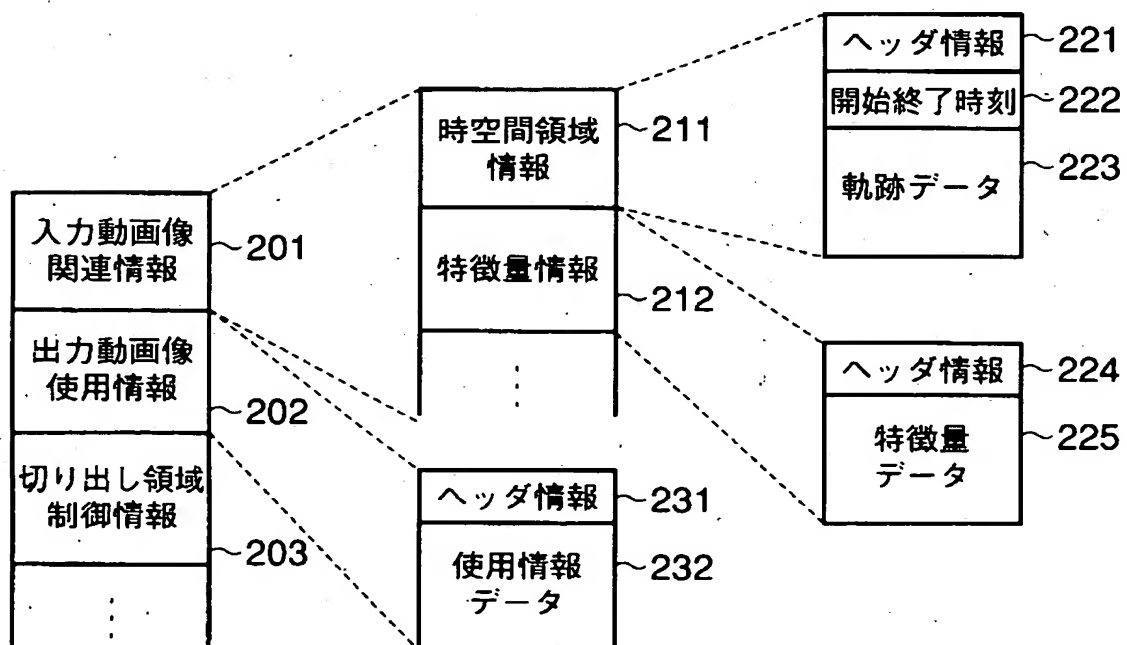
【書類名】

図面

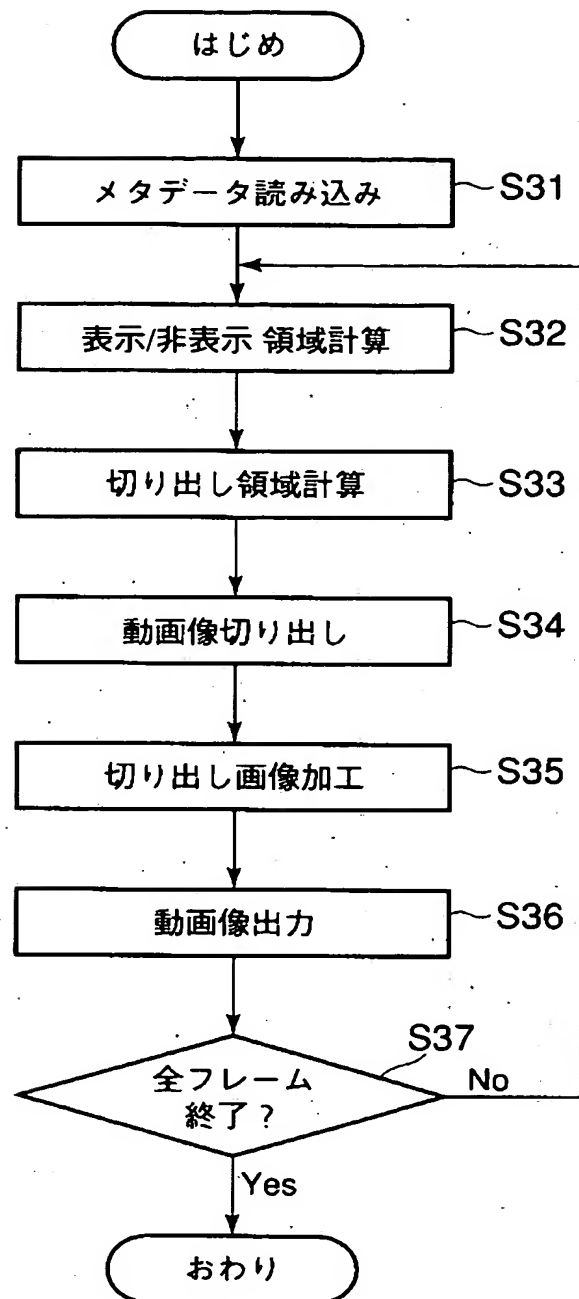
【図 1】



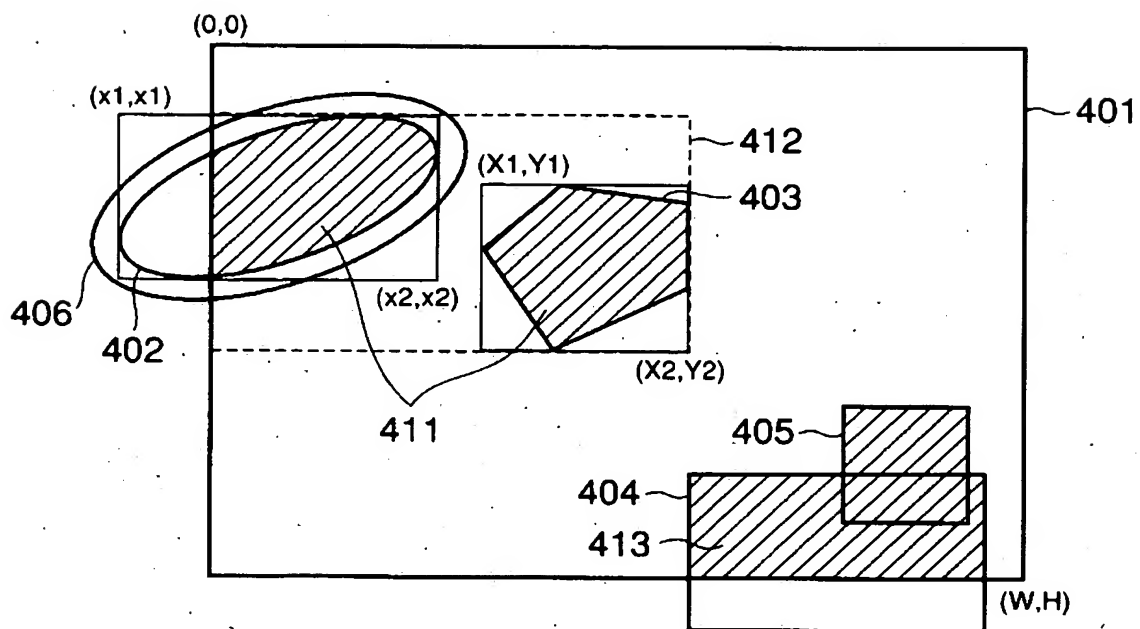
【図 2】



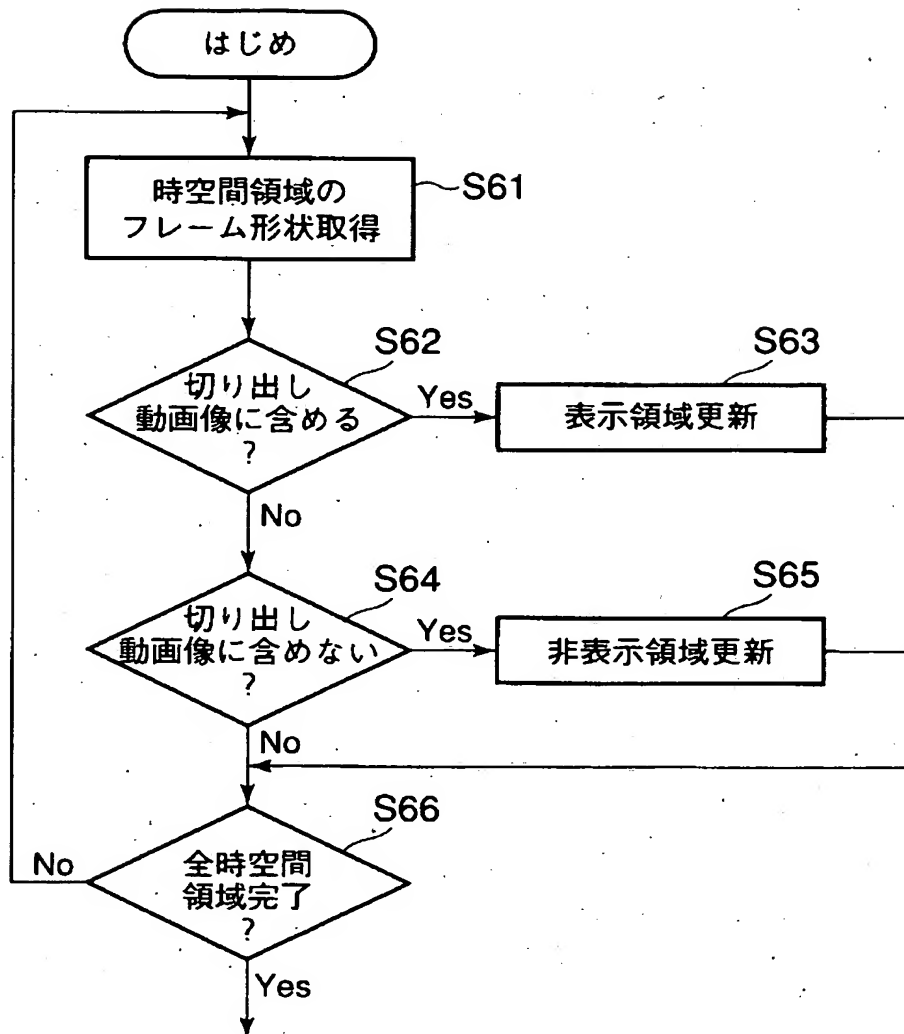
【図 3】



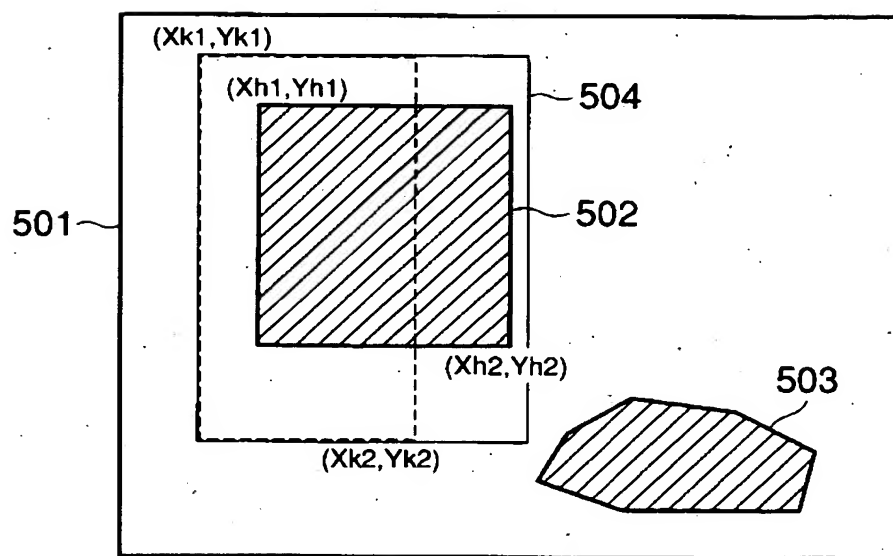
【図 4】



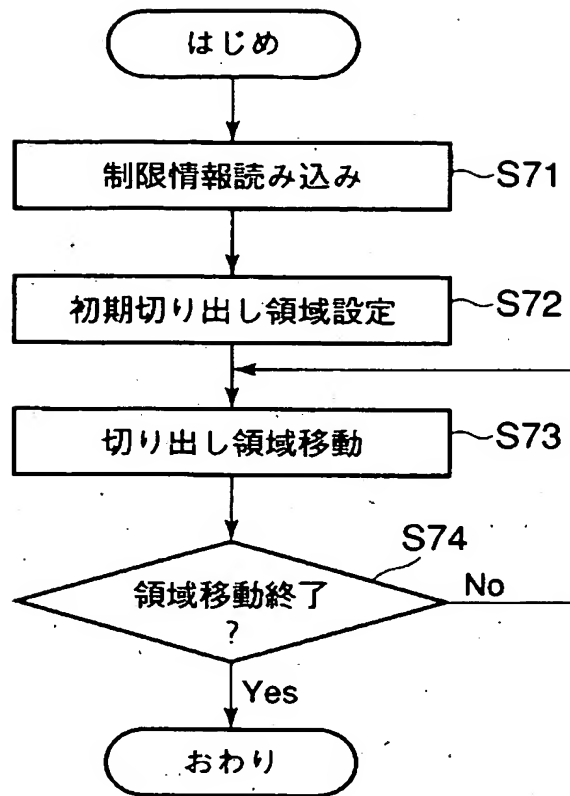
【図 5】



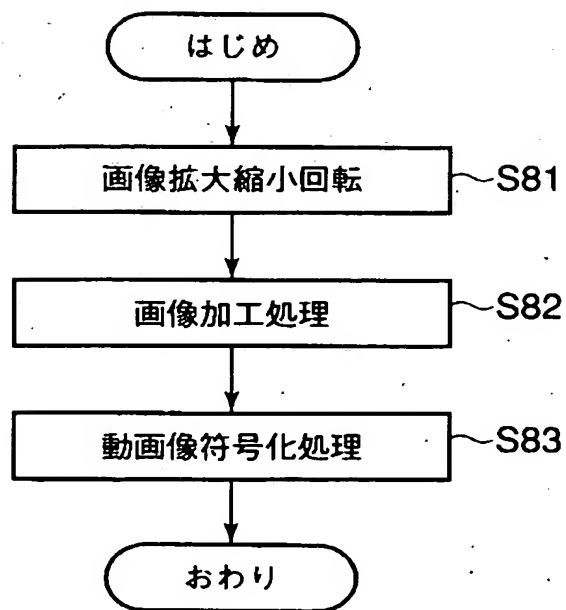
【図 6】



【図 7】



【図 8】





【書類名】 要約書

【要約】

【課題】 付帯するメタデータを利用して入力動画像の各フレームごとに適切な領域で切り出すことによって動画像特徴量や使用方法に応じた出力動画像を自動的に作成する動画像加工方法及び動画像加工装置を提供する。

【解決手段】 入力動画像を記憶する入力動画像記憶部 101 と、メタデータを記憶するメタデータ記憶部 102 と、各フレームにおける矩形の切り出し領域を決定する切り出し領域決定部 103 と、入力動画像の各フレームの画像から切り出し領域部の画像を切り出して出力動画像を作成する動画像切り出し部 104 と、出力動画像を表示する出力動画像表示部 105 と、出力動画像を記憶する出力動画像記憶 106 とを備える。

【選択図】 図 1

特願 2002-332756

出願人履歴情報

識別番号

[000003078]

1. 変更年月日 2001年 7月 2日  
[変更理由] 住所変更  
住 所 東京都港区芝浦一丁目1番1号  
氏 名 株式会社東芝
2. 変更年月日 2003年 5月 9日  
[変更理由] 名称変更  
住所変更  
住 所 東京都港区芝浦一丁目1番1号  
氏 名 株式会社東芝